Capstone Proposal Template

**[project name]**

**Business Understanding**

* What problem are you trying to solve, or what question are you trying to answer?
* Why this topic? AKA why do you care, and why should anyone else?
* What industry/realm/domain does this apply to?
* Who is your target audience?
* What impact would your answer/solution have on the real world, if your analysis were to be used/put into production?
* What pre-existing projects/research/papers in this field have you explored, or what domain knowledge are you relying upon?
* The motivation for the project is described. (Saying you needed to do a capstone project for flatiron is not an appropriate motivation)

**Data Understanding**

* What data will you collect?
* Where did your raw data come from?
* Is there a plan for how to get the data?
* Are the features that will be used described clearly?
* Has anyone else worked on this specific problem/dataset? If so, how will your work build on theirs?

**Data Preparation**

* In what form is the data stored?
* What data types are the variables? Do you have any frequency counts or descriptive statistics yet?
* What kind of preprocessing steps do you foresee?
* What are some of the cleaning/pre-processing challenges for this data?
* What is the minimum number of rows you will have? (this can be a ballpark estimate)
* How are you planning to visualize the important aspects of this data to bring it to life?

**Modeling**

* What modeling techniques are most appropriate for your problem?
* What is your target variable? (remember - we require that you answer/solve a supervised problem for the capstone, thus you will need a target)
* What model are you planning to use as a baseline?
* Is this a regression or classification problem?

**Evaluation**

* What metrics will you use to determine success?
* What does the minimum viable product (MVP) involve? What is the smaller project that you can accomplish in a week that your overall project is based on?
* What are your level-up stretch goals? How will you improve your project between MVP and presentation?

**Deployment**

* Is the method for reporting final results described?
* Is there a plan for deployment? (web app)
* What is the functionality?

**Tools/Methodologies**

* What are some of the Python libraries you are planning to use to gather, clean, explore, and model your data?
* What modeling algorithms are you planning to use?
* Where will you be performing your analysis - on your machine or in the cloud?
* Will your data be stored on your machine or in the cloud?